

# QUEEN MARY, UNIVERSITY OF LONDON

MTH 4106

Introduction to Statistics

Practical 2

17 January 2012

---

If you didn't do so last week, please complete Practical 1 before going on to do the following.

Remember to move to the `Mathematics department >> Statistics I` environment!

Today we will

- retrieve last week's project,
- learn how to do a stem-and-leaf plot in Minitab,
- use the Help to find out how Minitab calculates certain descriptive statistics.

**1 (Retrieving a previous project)** Start Minitab and open the Minitab project that you saved in Practical 1. You can do this by opening the **Project Manager**, right-clicking on the folder labelled "Untitled", and choosing **Open Project**. You should find your saved project in your folder **STATS**. I called it `Poplar1.MPJ` but you may have called it something different.

How many graphs are open in your retrieved project? Count those that are minimized and any that are hidden behind the **Session window** and **Data window** as well as those that are obvious. Open graphs take up a lot of working memory. If you have not done so already, copy any graphs that you want into the Report Pad, save the Report Pad, and then close all the graphs.

Is the Report Pad still open? If not, that is fine. If it is open, then make sure that you have saved it (see last week's notes). Then either delete the contents of the Report Pad, or right-click on `ReportPad`, and choose **Move to Word Processor**, giving it a suitable name; this saves the old report and clears space to put things in a new project report. The result of this *may* be that you have two copies of last week's **Project Report** in your **STATS** folder: check this later, and delete one of the copies, if necessary, to save space.

**2 (Stem-and-leaf plots)** You can do a stem-and-leaf plot by

`Stat` → `EDA` ► `Stem-and-leaf...`

Enter `Diameter` in `Graph Variables`. Click in the `By variable` box and then enter `Site`. Click `OK`.

In the **Session Window** two stem-and-leaf plots are displayed. They are both for the variable `Diameter`, one for each value of `Site`.

In what ways are these stem-and-leaf plots different from those in lectures?

If it is not already open, open a report pad (you learnt how to do this in Practical 1). Paste both stem-and-leaf plots into your report. At the bottom of your report, write your comments about why these plots differ from those in lectures. Save the report in your STATS folder.

You may find it helpful to print the report and keep it in a ring-binder along with all your practical worksheets and notes.

**3 (Median and quartiles)** For each site, the stem-and-leaf plot of **Diameter** should enable you to easily put the data in order. For each site, use the formulas given in lectures to work out the median, lower quartile and upper quartile. Write them below.

In last week's practical, you used Minitab to find the median, lower quartile and upper quartile for **Diameter** at each site. These should be in the report that you wrote then, or may still be in your **Session Window**. Write them below.

**4 (Using Help to find out about descriptive statistics)** You should have found that Minitab agrees with us about the median but not about the quartiles. How can we find out how Minitab calculates the quartiles? What about other things that we have not met yet, like the trimmed mean?

Fortunately, Minitab has extensive help facilities.

Use

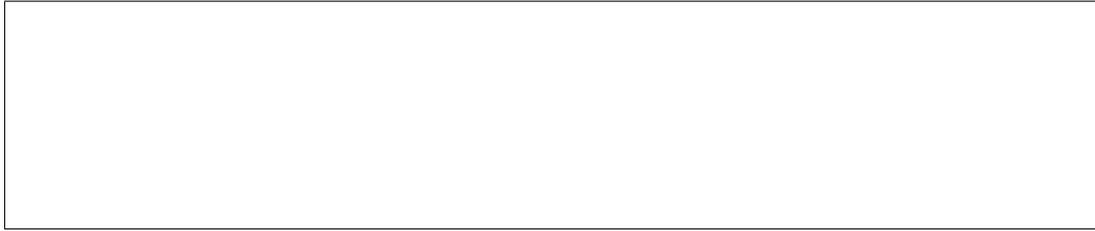
**Stat** → **Basic Statistics** ► **Display Descriptive Statistics**.

Now click on **Help** in the dialogue box. What comes up is general information about this command. We want to find out about some of the statistics so click on **Descriptive Statistics Available for Display or Storage**.

You will see a list of different statistics. Click on **Mean**, **Standard deviation** and **Trimmed Mean** and read what is displayed. Make notes about these below, or in a report pad. What are they used for? How are they calculated?

Similarly, click on **Median** and **First and third quartiles** to see what Minitab says about these quantities. Make notes about these below, or in a report pad.

Note that although a clear explanation of how the median is calculated is given, no exact details are given for the quartiles. Go back to the first **Help** window, and click on **see also** at the right-hand end of the top row, then on **Methods and formulas**. Find out exactly what formulas Minitab uses to calculate the sample variance, first quartile and third quartile.



I think that we can assume that Minitab's quartiles are close enough to ours for all practical purposes. I will not expect you to remember the formula which Minitab uses.

In any remaining time you should start work on the following question from Assignment 2. You can ask about it if necessary. Before starting this question, exit from this Minitab session and start again. Do not forget to save your project first! Note that you do not need to hand this assignment in until next week.

**5 (Assignment question)** To do this question from Assignment 2 you should open a new project. The question is copied here.

The following data set gives the height in centimetres of thirty female athletes. The data were collected at the Australian Institute of Sport, courtesy of Richard Telford and Ross Cunningham. Heights above the line correspond to basketball players, those below the line to rowers.

195.9	189.7	177.8	185.0	184.6	174.0
186.2	173.8	171.4	179.9	193.4	188.7
169.1					
177.9	177.5	179.6	181.3	179.7	185.2
177.3	179.3	175.3	174.0	183.3	184.7
180.2	180.7	176.0	156.0	179.7	

Use Minitab to write a short report including the following.

- Draw two boxplots, one for each sport.
- Draw two stem-and-leaf plots, one for each sport.
- Comment on the similarities and differences in the heights of the two groups of athletes which your plots have shown.
- Calculate the mean, median, minimum, maximum, lower and upper quartile, and standard deviation for each group of athletes. Do these statistics confirm your comments in part (c)?

Write your report in such a way that it is informative to the manager of the university sports centre. In particular, this means that diagrams should be clearly labelled, units of measurement should be shown, and comments should be set in the context of what the data are about.